ORIGINAL ARTICLE

Geospatial Artificial Intelligence: Potentials of Machine Learning for 3D Point Clouds and Geospatial Digital Twins

Jürgen Döllner[1]

Abstract

Artificial intelligence (AI) is changing fundamentally the way how IT solutions are implemented and operated across all application domains, including the geospatial domain. This contribution outlines AI-based techniques for 3D point clouds and geospatial digital twins as generic components of geospatial AI. First, we briefly reflect on the term "AI" and outline technology developments needed to apply AI to IT solutions, seen from a software engineering perspective. Next, we characterize 3D point clouds as key category of geodata and their role for creating the basis for geospatial digital twins; we explain the feasibility of machine learning (ML) and deep learning (DL) approaches for 3D point clouds. In particular, we argue that 3D point clouds can be seen as a corpus with similar properties as natural language corpora and formulate a "Naturalness Hypothesis" for 3D point clouds. In the main part, we introduce a workflow for interpreting 3D point clouds based on ML/DL approaches that derive domain-specific and application-specific semantics for 3D point clouds without having to create explicit spatial 3D models or explicit rule sets. Finally, examples are shown how ML/DL enables us to efficiently build and maintain base data for geospatial digital twins such as virtual 3D city models, indoor models, or building information models.

Keywords Geospatial artificial intelligence · Machine learning · Deep learning · 3D point clouds · Geospatial digital twins · 3D city models

Zusammenfassung

Georäumliche Künstliche Intelligenz: Potentiale des Maschinellen Lernens für 3D-Punktwolken und georäumliche digitale Zwillinge. Künstliche Intelligenz (KI) verändert grundlegend die Art und Weise, wie IT-Lösungen in allen Anwendungsbereichen, einschließlich dem Geoinformationsbereich, implementiert und betrieben werden. In diesem Beitrag stellen wir KI-basierte Techniken für 3D-Punktwolken als einen Baustein der georäumlichen KI vor. Zunächst werden kurz der Begriff „KI" und die technologischen Entwicklungen skizziert, die für die Anwendung von KI auf IT-Lösungen aus der Sicht der Softwaretechnik erforderlich sind. Als nächstes charakterisieren wir 3D-Punktwolken als Schlüsselkategorie von Geodaten und ihre Rolle für den Aufbau von räumlichen digitalen Zwillingen; wir erläutern die Machbarkeit der Ansätze für Maschinelles Lernen (ML) und Deep Learning (DL) in Bezug auf 3D-Punktwolken. Insbesondere argumentieren wir, dass 3D-Punktwolken als Korpus mit ähnlichen Eigenschaften wie natürlichsprachliche Korpusse gesehen werden können und formulieren eine „Natürlichkeitshypothese" für 3D-Punktwolken. Im Hauptteil stellen wir einen Workflow zur Interpretation von 3D-Punktwolken auf der Grundlage von ML/DL-Ansätzen vor, die eine domänenspezifische und anwendungsspezifische Semantik für 3D-Punktwolken ableiten, ohne explizite räumliche 3D-Modelle oder explizite Regelsätze erstellen zu müssen. Abschließend wird an Beispielen gezeigt, wie ML/DL es ermöglichen, Basisdaten für räumliche digitale Zwillinge, wie z.B. für virtuelle 3D-Stadtmodelle, Innenraummodelle oder Gebäudeinformationsmodelle, effizient aufzubauen und zu pflegen.

✉ Jürgen Döllner
doellner@hpi.uni-potsdam.de

1   Hasso Plattner Institute, Digital Engineering Faculty, University of Potsdam, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany

1 Introduction

Artificial intelligence (AI) is changing the way IT solutions are designed, built and operated. AI is not limited to specific application areas—it is finding its way into almost all industries and domains. As a collection of general-purpose

technologies, AI has significant impact because it "can transform opportunities not only for economic growth, but also for corporate profitability" (Purdy and Daugherty 2017).

For geospatial domains, fundamental questions include how AI can be specifically applied to or has to be specifically created for spatial data. Janowicz et al. (2020) give an overview of spatially explicit AI, which "utilizes advancements in techniques and data cultures to support the creation of more intelligent geographic information as well as methods, systems, and services for a variety of downstream tasks". This arising scientific discipline, called geospatial artificial intelligence (GeoAI), which "combines innovations in spatial science, artificial intelligence methods in machine learning (e.g. deep learning), data mining, and high-performance computing to extract knowledge from spatial big data" (Vopham et al. 2018), will in particular improve existing and create new technologies for geospatial information systems (GIS).

The relevance of AI for geospatial domains has been realized many years ago already regarding, for example, expert systems and knowledge-based systems (Openshaw and Openshaw 1997), geographically problem solving (Smith 1984), or analysing social sensing data (Wang et al. 2018a). In this paper, we concentrate on AI-based approaches for a specific category of 3D geodata, 3D point clouds, which are fundamental in photogrammetry, remote sensing and computer vision (Weinmann et al. 2015) and have manifold applications for building geospatial digital twins.

1.1 The Term "Artificial Intelligence"

The notion "AI" implies a number of conceptual difficulties, such as the definition of "natural", "human" or "general-purpose" intelligence. Put simply, the "most common misconception about artificial intelligence begins with the common misconception about natural intelligence. This misconception is that intelligence is a single dimension" (Kelly 2017). In the general public, AI often triggers associations and expectations such as simulating or overcoming human intelligence. If AI is pragmatically seen as technological progress, then "AI is going to amplify human intelligence not replace it, the same way any tool amplifies our abilities" (Lecun 2017). One of the AI applications that exemplified these controversies was ELIZA. This famous first chatbot, built by Josef Weizenbaum in 1966 (Weizenbaum 1966), was a speech-based simulation of a psychologist's interaction with a patient, which "demonstrated the kind of risk potential what was enclosed within such technological developments" (Palatini 2014). This topic is discussed further by Copeland (1993), who analyses which challenges and obstacles AI needs be to solved before "thinking" machines could be constructed. As he states, the key to AI is the ability of computers to think rationally, to discover meaning, to

generalize and learn from past experiences, and to be intelligent using learning, thinking, problem solving, perception and language.

There is generally no sharp border between AI and non-AI technology. Consider, for example, an autopilot steering an aircraft: in its beginning, it was perceived as AI, while today it has become a common operating technical component. That is to say, "when AI reaches mainstream usage it is frequently no longer considered as such" (Haenlein and Kaplan 2019). In that sense, the term "AI" is mostly used to label such technology that goes beyond current technology boundaries. In this contribution, we refer to AI in a geospatial context and focus on potentials of ML and DL for 3D point clouds.

1.2 AI-Based IT

From a software engineering perspective, several technologies and disciplines (Fig. 1) are required for the implementation of AI-based IT solutions:

- *Big data management* Many AI-based techniques require big data to be applied effectively. For example, ML and DL need training data, which is typically distilled from big data. Big data generally are characterized by a number of key traits (Kitchin and McArdle 2016), including:

  - large data amounts (volume),
  - rapid data capturing or generation (velocity),
  - different data types and structures (variety),
  - manifold relations among data sets (complexity),
  - high inherent data uncertainty (veracity).

  Big data have turned out to be a key driver for digital transformation processes as summarized in the famous statement "Data is the new oil. Data is just like crude.
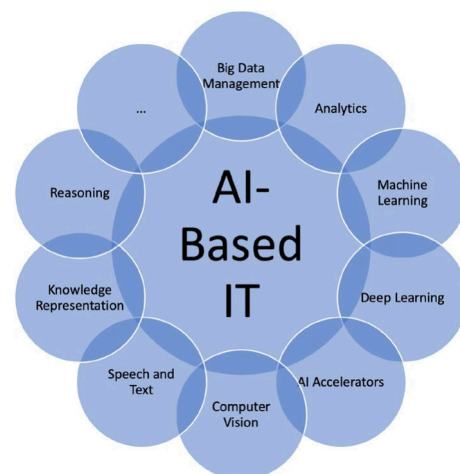

Fig. 1 Disciplines and technologies used to implement AI-based IT

It's valuable, but if unrefined it cannot really be used" (Humby 2006), which also is subject to a controversy (Marr 2018).

A broad range of approaches exist to capture, synthesize and simulate data about our geospatial reality. In that respect, geospatial data generally represent big data and the "oil" for the geospatial digital economy.

- *Analytics* Analytics refers to analytical reasoning and aims at providing concepts, methods, techniques, and tools to efficiently collect, organize, and analyse big data. Its objectives include to examine data, to draw conclusions, to get insights, to acquire knowledge, and to support decision making. In that respect, analytics "can be viewed as a sub-process in the overall process of 'insight extraction' from big data" (Gandomi and Haider 2015). For all variants, such as descriptive, predictive, and prescriptive analytics, ML/DL approaches support "the analysis and learning of massive amounts of unsupervised data, making it a valuable tool for big data analytics where raw data are largely unlabeled and uncategorized" (Najafabadi et al. 2015).

- *Machine learning* "Machine learning is programming computers to optimize a performance criterion using example data or past experience", which is required, in particular, "where we cannot write directly a computer program to solve a given problem" (Alpaydin 2014). That is to say, IT solutions do not rely on explicit or procedural problem solving strategies but are based on processing and analysing patterns and inference. For that reason, ML offers a different programming paradigm for the implementation of geospatial IT solutions.

  ML techniques can be generally classified into supervised ML, unsupervised ML, and reinforcement learning. Supervised ML, for example, acquires knowledge by building mathematical models based on training data, which are used to predict labels for input data. The underlying models "may be predictive to make predictions in the future, or descriptive to gain knowledge from data, or both" (Alpaydin 2014):

  "The promise and power of machine learning rest on its ability to generalize from examples and to handle noise" (Allamanis et al. 2018). For this purpose, ML offers a high degree of robustness regarding the input data set. A fundamental risk, however, lies in "overtraining" the model. If overtrained, it will perform very well on the training data, but will poorly generalize to new data. The model becomes incapable of generalizing, i.e. it is overfitting the training data. For that reason, "properly controlling or regularizing the training is key to out-of-sample generalization" (Zhang et al. 2018).

- *Deep learning* As Hatcher and Yu point out, DL applies "multi-neuron, multi-layer neural networks to perform learning tasks, including regression, classification, clustering, auto-encoding, and others" (Hatcher and Yu 2018). As a specific form of representation learning, which in turn is a specific form of ML, DL is based on artificial neural networks (ANNs) such as convolutional neural networks (CNNs) (Goodfellow et al. 2016). In the end, DL represents "essentially a statistical technique for classifying patterns, based on sample data, using neural networks with multiple layers" (Marcus 2018).

  It builds representations expressed in terms of simpler representations, i.e. we can build complex concepts out of simpler concepts. "It has turned out to be very good at discovering intricate structures in high-dimensional data and is therefore applicable to many domains of science, business and government" (LeCun et al. 2015), but is also faced by a number of limitations such as being "data hungry" and low support for transfer or hierarchical data (Marcus 2018).

- *AI accelerators* The growth of AI-based applications comes along with commodity, cost-efficient graphics processing units (GPUs) that are evolving to become high-performance accelerators for data-parallel computing. Further, there is a growing number of purpose-built systems for DL, for example, tensor processing units (TPUs), along with fully integrated hardware and software [e.g. NVidia DGX/HGX; TensorFlow (Abadi et al. 2016)]. They are known as AI accelerators, which differ with respect to hardware costs, training performance, computing power, and energy consumption (Wang et al. 2019) but foster and simplify AI-based software development and operation (Reuther et al. 2019).

To implement AI-based IT, components also include frameworks and systems for computer vision (e.g. image analysis, image understanding), speech and text analysis (e.g. text-to-speech, speech-to-text), knowledge representation and discovery, reasoning systems, etc.

1.3 Feature Space

ML/DL-based solutions describe the input data by means of feature vectors, i.e. by n-dimensional vectors whose numerical components describe selected aspects of a phenomenon to be observed, that is, the feature space equals a high-dimensional vector space. Dimensionality reduction techniques allow us to transform "high-dimensional data into a meaningful representation of reduced dimensionality" (Van Der Maaten et al. 2009) and, this way, to efficiently manage, process, and visualize feature spaces.

One of the key concepts required for geospatial ML and DL consists in finding adequate feature spaces for geospatial entities. A single point of the 3D point cloud does not allow for constructing meaningful feature vectors. To do so,

the local neighbourhood of the point must be analysed by selecting, for example, the k nearest points or points within a radius r (Weinmann et al. 2015). Next, geometric, topologic or any other high-level features are computed from the neighbourhood region that constitute components of the feature vector. "Typical recurring features include computing planarity, linearity, scatter, surface variance, vertical range, point colour, eigenentropy and omnivariance" (Griffiths and Boehm 2019).

2 3D Point Clouds

As universal 3D representations, 3D point clouds "can represent almost any type of physical object, site, landscape, geographic region, or infrastructure—at all scales and with any precision" as Richter (2018) states, who discusses algorithms and data structures for out-of-core processing, analysing, and classifying of 3D point clouds. To acquire 3D point clouds, various technologies can be applied including airborne or terrestrial laser scanning, mobile mapping, RGB-D cameras (Zollhöfer et al. 2018), image matching, or multi-beam echo sounding.

2.1 Geospatial Digital Twins

3D point clouds are ubiquitous for geospatial applications such as for environmental monitoring, disaster management, urban planning, building information models, or self-driving vehicles. More precisely, 3D point clouds are commonly used as base data for reconstructing 3D models (e.g. digital terrain models, virtual 3D city models, building information models), but can also be understood as point-based 3D models, for example, in the case when 3D point clouds are dense (Fig. 2).

In particular, they represent key components of geospatial digital twins, that is, digital replica of spatial entities and phenomena. Digital twins, in general, are composed of three parts, "which are the physical entities in the physical world, the virtual models in the virtual world, and the connected data that tie the two worlds" (Qi and Tao 2018). While the connection between both can be handled by sensors, the virtual models have to be derived from the physical counterparts. For example, 3D point clouds are used to derive 3D indoor models, which are essential components for real-time building information models together with sensor networks and IoT devices (Khajavi et al. 2019); they also "represent a generic approach to capture, model, analyse, and visualize digital twins used by operators in Industry 4.0 application scenarios" (Posada et al. 2018). In that regard, geospatial digital twins are means for monitoring, visualizing, exploring, optimizing,

Fig. 2 Example of a high-density 3D point cloud of an indoor environment

and predicting behaviour and processes related to the corresponding physical entities.

2.2 Characteristics

A 3D point cloud represents a set of 3D points in a given coordinate system and can be characterized by:

- *Uniform representation*—unstructured, unordered set of 3D points (e.g. in an Euclidian space);
- *Discrete representation*—discrete samples of shapes without restrictions regarding topology or geometry;
- *Irregularity*—expose irregular spatial distribution and varying spatial density;
- *Incompleteness*—due to the discrete sampling, representations are incomplete by nature;
- *Ambiguity*—the semantics (e.g. surface type, object type) of a single point generally cannot be determined without considering its neighbourhood;
- *Per-point attributes*—each point can be attributed by additional per-point data such as colour or surface normal;
- *Massiveness*—depending on the density of the capturing technology, 3D point clouds may consist of millions or billions of points.

The key trait of 3D point clouds, a fundamental category of 3D geospatial data, consists in the absence of any structural, hierarchical, or semantics-related information—a 3D point cloud is a simple, unordered set of 3D points. "The lack of topology and connectivity, however, is strength and weakness at the same time" (Gross and Pfister 2007). There is, therefore, a strong demand for solutions that allow us to enrich 3D point clouds with information.

2.3 3D Point Cloud Time Series

For a growing number of applications, 3D point clouds are captured and processed with high frequency. For example, if

a surveillance system captures its target environment every second, it results in a stream of 3D point clouds .

If 3D point clouds are captured or generated at different points in time having overlapping geospatial regions, these sets are inherently related. By 3D point cloud time series, we refer to a collection of 3D point clouds taken at different points in time for a common geospatial region. The collection of 3D point clouds represents, in a sense, a 4D point cloud.

3D point cloud time series have a high degree of redundancy, which needs to be exploited to achieve efficient management, processing, compression and storage, for example, separating static from dynamic structures. Redundancy can also be used to improve accuracy and robustness of 3D point cloud interpretations and related predictions.

2.4 Feasibility of ML/DL-Based Approaches

Kanevski et al. investigate the general applicability of ML to geospatial data and conclude "the key feature of the ML models/algorithms is that they learn from data and can be used in cases when the modelled phenomenon is not very well described, which is the case in many applications of geospatial data" (Kanevski et al. 2009). The complete absence of structure, order and semantics as well as the inherent irregularity, incompleteness, and ambiguity explain why 3D point clouds are not very well described and difficult candidates for procedural and algorithmic programming. Their characteristics, however, allow us to effectively apply ML/DL to 3D point clouds:

- *Big data*: 3D point clouds are spatial big data that can be cost-efficiently generated for almost all types of spatial environments—big data are a prerequisite for ML/DL-based approaches;
- *Fuzziness*: 3D point clouds show inherent fuzziness and noise as they are sampling shapes by means of discrete representations—ML/DL are particularly handling well fuzzy and noise data;
- *Semantics*: Depending on the concrete application domain, semantic concepts can be defined and corresponding training data can be configured to distill the required semantics.

In the case of 3D point clouds, ML/DL support computing domain-specific and application-specific information, typically by point classification, point cloud segmentation, object identification, and shape reconstruction. Compared to traditional procedural-like, heuristic, or empiric-based algorithms, ML/DL-based techniques generally have significantly less implementation complexity as the implementation relies on general-purpose AI frameworks having a high robustness and high rate of innovation. The customization

consists in finding appropriate training and test data. A crucial issue represents robustness of ML techniques as ML models "are vulnerable to adversarial examples formed by applying small carefully chosen perturbations to inputs that cause unexpected classification errors" (Rozsa et al. 2016).

2.5 Naturalness Hypothesis

The "Naturalness Hypothesis" (Allamanis et al. 2018), which is investigated in natural language recognition and software analytics, helps understanding further why ML and DL approaches provide effective instruments for analysing and interpreting 3D point clouds. In general, one key approach to ML and DL is to find out whether a given problem domain corresponds to or has similar statistical properties as large natural language corpora (Jurafsky and Martin 2000). Here, ML/DL-based approaches have shown extraordinary success in natural language recognition, natural language translation, question answering, text mining, text comprehension, etc. The most important finding in these areas is that objects (e.g. spoken or written texts) are less diverse than they initially seem: most human expressions ("utterances") are much simpler, much more repetitive, and much more predictable than the expressiveness of the language body suggests, whereby "these utterances can be very usefully modelled using modern statistical methods" (Hindle et al. 2012). This phenomenon can be understood with measures of perplexity and cross-entropy (de Boer et al. 2005).

3D point clouds seen as a form of natural communication as well as geospatial environments are ultimately repetitive regardless of the endless variations they may exhibit. In a sense, 3D point clouds are just "spatial utterances" that can be modelled using statistical methods. 3D point clouds, thereby, constitute 3D point cloud corpora to which ML technology can be applied taking advantage of the statistical distributional properties estimated over representative point cloud corpora.

Following the schema for an argumentation originally set up for software engineering (Hindle et al. 2012), we formulate an ML/DL-centric naturalness hypothesis for 3D point clouds as follows:

> 3D point clouds, in theory, are complex, expressive and powerful, but the 3D point clouds actually captured or generated in geospatial domains are far less complex, far less expressive and repetitive. Their predictable statistical properties can be captured in statistical language models and leveraged for geospatial data analysis.

General ML and DL approaches, however, need to be adapted to the characteristics of 3D point clouds. "Most critically, standard deep neural network models require input

data with regular structure, while point clouds are fundamentally irregular: point positions are continuously distributed in the space and any permutation of their ordering does not change the spatial distribution" (Wang et al. 2018b).

3 ML/DL-Based Point Cloud Interpretation

3D point clouds provide cost-efficient raw data for creating the basis for geospatial digital twins at all scales but they are purely geometric data without any structural or semantics information about the objects they represent. Motivated by the naturalness hypothesis, ML and DL can be applied to analyse and interpret that data as well as to provide powerful techniques if it comes to discrete irregular, incomplete, and ambiguous data of a given corpus—exactly what characterizes 3D point clouds.

3.1 Interpretation Concept

The ML/DL-based processing of 3D point clouds is based on the concept of interpretation known from programming languages. In that scheme, however, analytics and semantics derivation do not require steps that "compile" raw data into higher-level representations. To process data, for example, the PointNet neural network "directly consumes point clouds and well respects the permutation invariance of points in the input" and provides a "unified architecture for applications ranging from object classification, part segmentation, to scene semantic parsing" (Qi et al. 2016a).

Applications or services that require spatial information encoded in 3D point clouds specify the exact set of features to be extracted and the spatial extent to be searched. The available feature types depend on how the ML and DL subsystems have been trained before. The analysis processes the request by triggering the evaluation to obtain the results. At no point in time, 3D point clouds are pre-processed or pre-evaluated nor do they require any intermediate representations, i.e. the interpretation works on-demand on the raw point cloud data.

In Fig. 3 the classical geoprocessing workflow (a) is compared to the workflow enabled by 3D point cloud interpretation (b). While (a) is based on generating more and more detailed and semantically well-defined representations, the workflow (b) operates on raw data, extracting the demanded features using corresponding training data. The ML/DL engine that implements that workflow needs core functionality such as:

- *Point classification*: According to defined point categories (e.g vegetation, built structures, water, streets) labels are computed and attached as per-point attributes together with the probability for this category assignment. For example, Roveri et al. "automatically transform the 3D unordered input data into a set of useful 2D depth
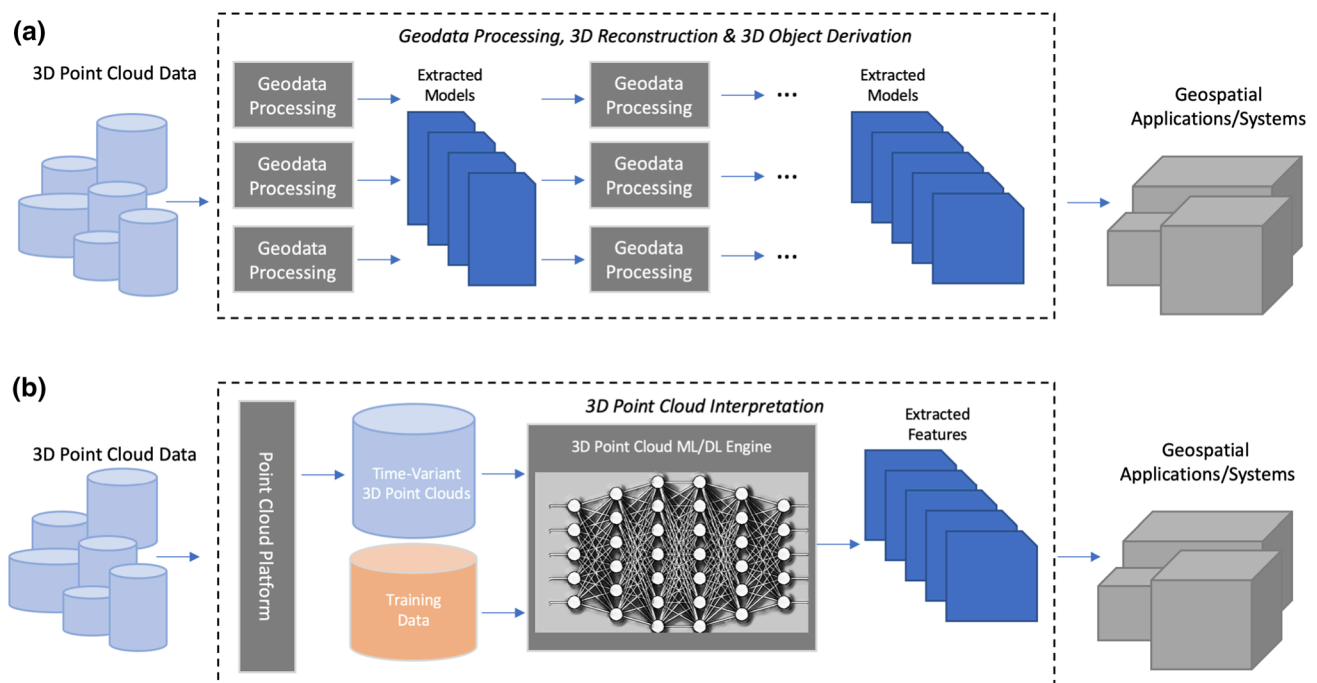

Fig. 3 3D point cloud processing: classical workflow **a** based on 3D reconstruction, 3D modelling, and object derivation; **b** ML/DL workflow based on 3D point cloud interpretation

images, and classify them by exploiting well-performing image classification CNNs" (Roveri et al. 2018).

- *Point cloud segmentation*: Segmentation as a core operation for 3D point clouds helps reducing fragmentation and subdividing large point clouds. Typically, it is based on identifying 3D geometry features such as edges, planar facets, or corners. ML and DL, in contrast, allow us to take advantage of semantic cues and affordances found in 3D point clouds. For example, we can segment "local geometric structures by constructing a local neighbourhood graph and applying convolution-like operations on the edges connecting neighbouring pairs of points, in the spirit of graph neural networks" (Wang et al. 2018b).

- *Shape recognition*: Shapes are essential for understanding 3D environments. To recognize them, a combined 2D–3D approach (Stojanovic et al. 2019b) consists of generating 2D renderings from 3D point clouds that are evaluated by image analysis. For this purpose, CNNs can combine "information from multiple views of a 3D shape into a single and compact shape descriptor offering even better recognition performance" (Su et al. 2015) compared to approaches that operate directly on raw 3D point clouds. Large general-purpose repositories of 3D objects, in addition, provide a solid training data base.

- *Object classification*: Applications generally require object-based information to be extracted from 3D point clouds, for example, signs and poles of the street space. Based on classified and segmented 3D point clouds, CNNs based upon volumetric representations or CNNs based upon multi-view representations are commonly applied to this end; Qi et al. (2016b) give an overview of the space of methods available.

The non-uniform sampling density typically found in 3D point clouds represents a key challenge for ML/DL-based learning. Qi et al. (2017) propose a hierarchical CNN that operates on nested partitions of an input point set because "features learned in dense data may not generalize to sparsely sampled regions. Consequently, models trained for sparse point cloud may not recognize fine-grained local structures".

ML/DL-based interpretation enables us to implement generic analysis components for 3D point clouds. As no intermediate representations are required, analysis results are only created once they are requested and they are only computed for the specific region the application has defined. Among the advantages of this approach are:

- *Configurability*: The ML/DL training data together with feature vector definitions allow for many label types to be predicted. For it, the generic, domain-independent mechanism offers a high degree of configurability.
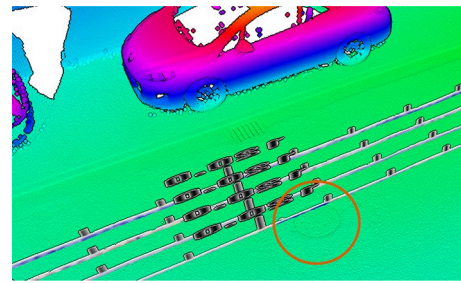

Fig. 4 Example of an analysis of underground structures based on 3D point clouds captured by radar combined with four trajectories of ground-penetrating radar data. The point cloud is coloured with a height gradient. Dataset from the city of Essen, Germany

- *Service-based computing*: The approach is scalable as it can be fully mapped to a service-oriented architecture and scalable hardware (e.g. GPU clusters), built by lower-level and higher-level services and mashups.

- *On-demand computation*: Downstream services allow for on-demand computation. For many classifications, the intepretation can be executed even in real time (e.g. object detection out of point clouds for surveillance purposes).

- *Raw data processing*: Storage and handling of massive 3D point clouds, including time-variant ones, can be optimized independently as the interpretation only requires fast spatial access to point cloud contents.

- *Storage efficiency*: There are no pre-selected or pre-built 3D models or intermediate representations. The approach therefore works well for massive or time-varying 3D point clouds. In particular, the original precision of the raw data is never reduced as raw data are fed directly into the ML/DL processes.

3.2 Examples

In a joint research project, we are developing a robust, high-performance engine for experimental ML/DL-based geospatial analytics. It provides features to store, manage, and visualize massive 4D point clouds.

In Fig. 4, 3D point cloud interpretation has been used to extract the underground infrastructure entities in the street space from mobile mapping data (Wolf et al. 2019). The visualization shows the extracted tubes and also has detected street elements such as manhole covers. In Fig. 5, 3D point cloud interpretation has identified and classified points according to different categories of street space furniture.

Figure 6 shows how time series of 3D point clouds, for example taken during a mobile scan, can be interpreted to extract relevant objects such as street signs, vehicles, vegetation, etc.

In Fig. 7, a composite classification is illustrated: the bike and the person riding the bike are identified and then can be combined as 'person-riding-a-bike'. High-level abstractions can be built in a post-processing step or as part of the ML/DL processes.

4 AI for Geospatial Digital Twins

A key demand in digital transformation processes represent digital twins in the sense of digital representations and replica that reflect key traits, behaviour and states of a living or non-living physical entity (El Saddik 2018). The construction of base data for geospatial digital twins based on explicitly defined 3D model schemata is a labour-intensive and error-prone process, for example, virtual 3D city models with high level of detail such as CityGML LOD3 or LOD4 (Löwner et al. 2016) as 3D reconstruction processes as well as the modelling schemata must deal with inaccurate, incomplete data and generally cannot deal with special cases that are not provided in the modelling scheme. Whether we apply strong mathematics or fine-tuned heuristics, a reconstructed 3D model almost always lacks details and it can hardly mirror weakly sampled, unusual, or fuzzy entities.

3D point clouds represent raw data of geospatial entities in a well-defined, consistent, and simple way, in particular, for spatial environments such as indoor spaces (Stojanovic et al. 2019a), building information models, and cities. ML/DL-based interpretation can both efficiently and effectively, analyse and organize 3D point clouds without being restricted by explicitly defined modelling schemata. Above all, it flexibly generates semantics on-demand and on-the-fly, that is, it helps "healing" one of the biggest weaknesses of 3D point clouds—the lack of structure and semantics. There is virtually no limitation for the specific types of 3D objects, structures, or phenomena that can be identified and extracted by ML/DL-based 3D point cloud interpretation.

In addition, ML/DL-based interpretation operates on the raw geospatial data, i.e. it retains a high degree of
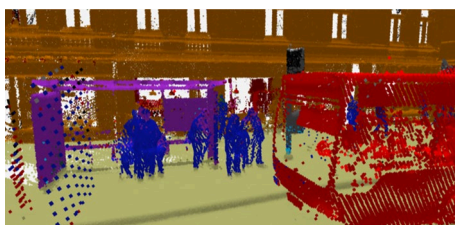

Fig. 6 Example of object classification (based on PointNet) within a dynamic scenario given by a time series of 3D point clouds

originality, while using only a moderate degree of explicit modelling for extracted features (Fig. 8). This, on the one hand, simplifies management and storage, in particular, if it comes to time series. On the other hand, it helps identifying and classifying ambiguous or fuzzy entities as needed to, for example robustly and automatically build geospatial digital twins.

5 Conclusions

AI is radically changing programming paradigms and software solutions in all application domains. In the geospatial domain, the data characteristics are particularly suitable for ML/DL approaches as geodata fits into the concept of a "linguistic corpus" as sketched in the context of the naturalness hypothesis. ML/DL-based analysis and extraction of features out of 3D point clouds, for example, can be used to derive application-specific, domain-specific and task-specific semantics.

Above all, ML/DL-based interpretation of 3D point clouds enables us to transcend explicit geospatial modelling and, therefore, to overcome complex, heuristics-based reconstructions and model-based abstractions. In that regard,

Fig. 5 Example of point classification for a typical street space scenario. Ground, buildings, vehicles, pedestrians, and different street furniture objects are classified with a PointNet-based approach and are visualized by different colours
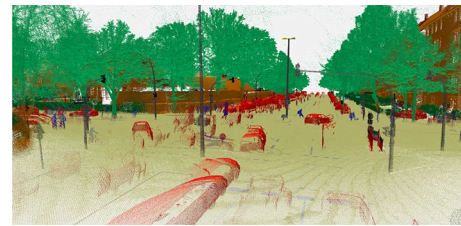

Fig. 7 Example of a composite classification (data by courtesy of Stadt Hamburg). Bicycle and cyclist are segregated by DL-based classification
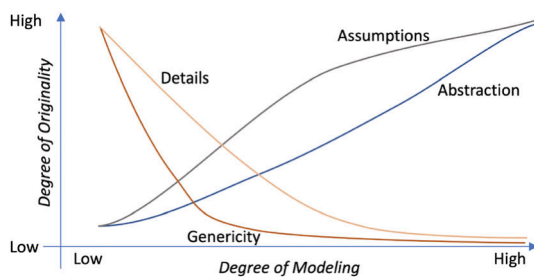
Fig. 8 The higher the degree of explicit modelling, the less of the original raw data is preserved as more and more assumptions and abstraction are introduced. Hence, data details disappear and less general-purpose models result

AI technology can be used to simplify and accelerate workflows for geodata processing and geoinformation systems. Of course, crucial ML/DL-related challenges result from the demand for effective training data and efficient feature representations.

Last but not least, ML/DL-based solutions offer simplifications in the dimension of software engineering. Large parts of today's implementations (often historically grown with large amounts of so called technical debts) will be partically replaced by ML/DL "black box" subsystems, which have far less software management and software development complexities. In particular, most heuristics-based, explicitly programmed analysis routines, which tend to be difficult to parameterize and to configure, can be migrated this way. As a consequence, ML/DL-based approaches, in the long run, have the potential to "eat up" many of today's explicitly programmed GIS implementations.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow IJ, Harp A, Irving G, Isard M (2016) TensorFlow: large-scale machine learning on heterogeneous distributed systems. CoRR arXiv:1603.04467

Allamanis M, Barr E, Devanbu P, Sutton C (2018) A survey of machine learning for big code and naturalness. ACM Comput Surv 51(4):81:1–81:37

Alpaydin E (2014) Introduction to machine learning, 3rd edn. MIT Press, Adaptive Computation and Machine Learning

Copeland J (1993) Artificial Intelligence: a philosophical introduction. Wiley-Blackwell, USA

de Boer P, Kroese D, Shie M, Rubinstein R (2005) A tutorial on the cross-entropy method. Ann Operations Res 134(1):19–67

El Saddik A (2018) Digital twins: the convergence of multimedia technologies. IEEE MultiMedia 25(2):87–92

Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag 35(2):137–144

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press. http://www.deeplearningbook.org

Griffiths D, Boehm J (2019) A review on deep learning techniques for 3D sensed data classification. CoRR arXiv:abs/1907.04444,

Gross M, Pfister H (2007) Point-based graphics. Morgan Kaufmann Publishers Inc, USA

Haenlein M, Kaplan A (2019) A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. Calif Manag Rev 61(4):5–14

Hatcher W, Yu W (2018) A survey of deep learning: platforms, applications and emerging research trends. IEEE Access 6:24411–24432. https://doi.org/10.1109/ACCESS.2018.2830661

Hindle A, Barr E, Su Z, Gabel M, Devanbu P (2012) On the Naturalness of Software. In: Proceedings of the 34th international conference on software engineering, IEEE Press, ICSE '12, pp 837–847

Humby C (2006) http://www.humbyanddunn.com

Janowicz K, Gao S, McKenzie G, Hu Y, Bhaduri B (2020) GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. Int J Geogr Inf Sci. https://doi.org/10.1080/13658816.2019.1684500

Jurafsky D, Martin J (2000) Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition, 1st edn. Prentice Hall PTR, USA

Kanevski M, Foresti L, Kaiser C, Pozdnoukhov A, Timonin V, Tuia D (2009) Machine learning models for geospatial data. Handbook of theoretical and quantitative geography. University of Lausanne, Lausanne, pp 175–227

Kelly K (2017) The AI Cargo Cult: the myth of a superhuman AI. Tech. rep., Backchannel. https://www.wired.com/2017/04/the-myth-of-a-superhuman-ai

Khajavi SH, Hossein Motlagh N, Jaribion A, Werner LC, Holmström J (2019) Digital twin: vision, benefits, boundaries, and creation for buildings. IEEE Access 7:147406–147419. https://doi.org/10.1109/ACCESS.2019.2946515

Kitchin R, McArdle G (2016) What makes big data, big data? Exploring the ontological characteristics of 26 datasets. Big Data Soc. https://doi.org/10.1177/2053951716631130

Lecun Y (2017) AI is going to amplify human intelligence not replace it. FAZ Netzwirtschaft. https://www.faz.net/-gqm-8yrxk

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444. https://doi.org/10.1038/nature14539

Löwner MO, Gröger G, Benner J, Biljecki F, Nagel C (2016) Proposal for a new LOD and multi-representation concept for CityGML. In: ISPRS annals of the photogrammetry, remote sensing and spatial information sciences, vol 4, pp 3–12

Marcus G (2018) Deep learning: a critical appraisal. CoRR arXiv:abs/1801.00631

Marr B (2018) Here's why data is not the new oil. Forbes. www.forbes.com

Najafabadi MM, Villanustre F, Khoshgoftaar TM et al (2015) Deep learning applications and challenges in big data analytics. J Big Data 2:1. https://doi.org/10.1186/s40537-014-0007-7

Openshaw S, Openshaw C (1997) Artificial intelligence in geography. Wiley, USA

Palatini K (2014) Joseph Weizenbaum, responsibility and humanoid robots. In: Funk M, Irrgang B (eds) Robotics in Germany and Japan. Philosophical and technical perspectives. Peter Lang, USA, pp 163–169

Posada J, Zorrilla M, Dominguez A, Simoes B, Eisert P, Stricker D, Rambach J, Döllner J, Guevara M (2018) Graphics and media technologies for operators in industry 4.0. IEEE Comput Graph Appl 38(5):119–132

Purdy M, Daugherty P (2017) How AI boosts industry profits and innovation. http://www.accenture.com

Qi C, Su H, Mo K, Guibas L (2016a) PointNet: deep learning on point sets for 3D classification and segmentation. CoRR arXiv:abs/1612.00593

Qi C, Su H, Nießner M, Dai A, Yan M, Guibas L (2016b) Volumetric and multi-view CNNs for object classification on 3D data. CoRR arXiv:1604.03265

Qi Q, Tao F (2018) Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. IEEE Access 6:3585–3593

Qi V, Yi L, Su H, Guibas L (2017) PointNet++: Deep hierarchical feature learning on point sets in a metric space. CoRR arXiv:1706.02413

Reuther A, Michaleas P, Jones M, Gadepally V, Samsi S, Kepner J (2019) Survey and benchmarking of machine learning accelerators. In: 2019 IEEE high performance extreme computing conference (HPEC), pp 1–9

Richter R (2018) Concepts and techniques for processing and rendering of massive 3D point clouds. PhD thesis, University of Potsdam, Faculty of Digital Engineering, Hasso Plattner Institute

Roveri R, Rahmann L, Öztireli C, Gross M (2018) A network architecture for point cloud classification via automatic depth images generation. In: 2018 IEEE conference on computer vision and pattern recognition, CVPR 2018, Salt Lake City, UT, USA, June 18–22, 2018, pp 4176–4184

Rozsa A, Günther M, Boult T (2016) Are accuracy and robustness correlated. In: 2016 15th IEEE international conference on machine learning and applications (ICMLA), pp 227–232

Smith T (1984) Artificial intelligence and its applicability to geographical problem solving. Prof Geogrh 36(2):147–158

Stojanovic V, Trapp M, Döllner J, Richter R (2019a) Classification of indoor point clouds using multiviews. In: The 24th international conference on 3D web technology, Web3D, Los Angeles, July 26-28, 2019, pp 1–9

Stojanovic V, Trapp M, Richter R, Döllner J (2019b) Generation of approximate 2D and 3D floor plans from 3D point clouds. In: Proceedings of the 14th international joint conference on computer vision, imaging and computer graphics theory and applications, VISIGRAPP 2019, Vol 1: GRAPP, pp 177–184

Su H, Maji S, Kalogerakis E, Learned-Miller E (2015) Multi-view convolutional neural networks for 3D shape recognition. In: Proceedings of the 2015 IEEE international conference on computer vision (ICCV), IEEE Computer Society, ICCV '15, pp 945–953

Van Der Maaten L, Postma E, Van den Herik J (2009) Dimensionality reduction: a comparative review. J Mach Learn Res 10:66–71

Vopham T, Hart J, Laden F, Chiang Y (2018) Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. Environ Health. https://doi.org/10.1186/s12940-018-0386-x

Wang D, Szymanski B, Abdelzaher T, Ji H, Kaplan L (2018a) The age of social sensing. CoRR arXiv:abs/1801.09116

Wang Y, Sun Y, Liu Z, Sarma S, Bronstein M, Solomon J (2018b) Dynamic graph CNN for learning on point clouds. CoRR arXiv:1801.07829

Wang Y, Wang Q, Shi S, He X, Tang Z, Zhao K, Chu X (2019) Benchmarking the performance and power of AI accelerators for AI training. arXiv:1909.06842

Weinmann M, Schmidt A, Mallet C, Hinz S, Rottensteiner F, Jutzi B (2015) Contextual classification of point cloud data by exploiting individual 3D neighborhoods. ISPRS Ann Photogramm Remote Sens Spatial Inf Sci II–3/W4:271–278

Weizenbaum J (1966) ELIZA—a computer program for the study of natural language communication between man and machine. Commun ACM 9(1):36–45

Wolf J, Richter R, Döllner J (2019) Techniques for automated classification and segregation of mobile mapping 3D point clouds. In: Proceedings of the 14th international joint conference on computer vision, imaging and computer graphics theory and applications, VISIGRAPP 2019, vol 1: GRAPP, pp 201–208

Zhang C, Vinyals O, Munos R, Bengio S (2018) A study on overfitting in deep reinforcement learning. CoRR arXiv:1804.06893

Zollhöfer M, Stotko P, Görlitz A, Theobalt C, Niessner M, Klein R, Kolb A (2018) State of the art on 3D reconstruction with RGB-D cameras. Comput Graph Forum 37:625–652